

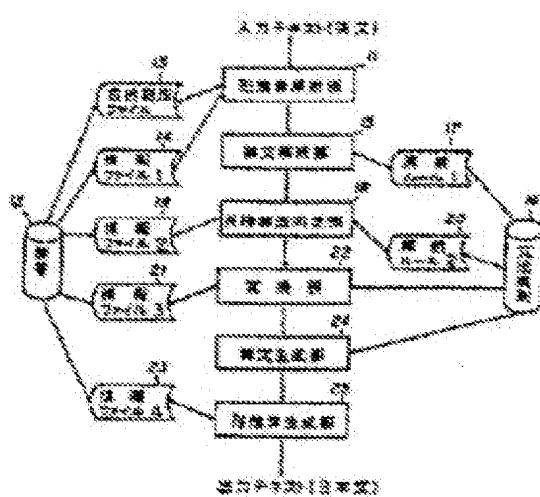
MACHINE TRANSLATION DEVICE

Publication number: JP4107769
Publication date: 1992-04-09
Inventor: OGURO YOSHIHISA
Applicant: RICOH KK
Classification:
- International: G06F17/28; G06F17/28; (IPC1-7): G06F15/38
- European:
Application number: JP19900227748 19900829
Priority number(s): JP19900227748 19900829

Report a data error here

Abstract of JP4107769

PURPOSE: To generate translated sentence, which is faithful to the conceptional structure of a source sentence and can be easily comprehended, by detecting indeclinable words, declinable words and case postpositional particles in the clause of the translated sentence and changing the order of words in the translated sentence according to the detected results. **CONSTITUTION:** This device is equipped with a morpheme analysis part 11, dictionary 12, inflective word form file 13, information file 14, syntax analysis part 15, grammatical rule 16, analysis rule 17, English construction identifying part 18, information file 19, analysis rule 20, information file 21, conversion part 22, information file 23, syntax generation part 24 and morpheme generation part 25. The source sentence of an objective language is translated into an objective language by using the analysis dictionary 12, and the inputted source sentence and the translated result are displayed. Then, declinable parts, declinable parts and case postpositional particles (corresponding phrases) in the clause of the translated sentence are detected and according to these detected results, the order of words in the translated sentence is changed. Thus, the natural and easily comprehensible translated sentence is generated without loosing the meaning of the conceptional structure of the original.



Data supplied from the esp@cenet database - Worldwide

◎公開特許公報(A)

平4-107769

◎Int.Cl. 3

G 06 F 15/38

識別記号

府内整理番号

T

9194-5L

◎公開 平成4年(1992)4月9日

審査請求 未請求 請求項の数 3 (全6頁)

◎発明の名称 機械翻訳装置

◎特 願 平2-227748

◎出 願 平2(1990)8月29日

◎発明者 大黒 要久 東京都大田区中馬込1丁目3番6号 株式会社リコー内

◎出願人 株式会社リコー 東京都大田区中馬込1丁目3番6号

◎代理人 弁理士 高野 明近 外1名

明細書

1. 発明の名称

機械翻訳装置

2. 技術請求の範囲

1. 少なくとも、対象言語の原文を入力するための入力手段と、該入力手段により入力された原文を認識する認識手段と、言語解析に必要な知識を記憶している解析装置と、該記憶手段に記憶された対象言語の原文を前記解析装置を用いて目的言語に翻訳する翻訳手段と、前記入力手段により入力された原文と前記翻訳手段により翻訳された翻訳結果を表示する表示手段と、原文の文節内の体言及び用言、格助詞を検出する検出手段と、該検出手段の検出結果によって該文の語彙を変更する語彙変更手段とを有することを特徴とする機械翻訳装置。

2. 前記語彙変更手段において、該語彙変更対象又は該語彙に勾配点を付与する付与手段を有することを特徴とする請求項1記載の機械翻訳装置。

3. 前記語彙変更手段において、前記検出手段

の検出結果だけでなく、原文側の語彙を反映した語彙変更手段を有することを特徴とする請求項1記載の機械翻訳装置。

3. 発明の詳細な説明

技術分類

本発明は、機械翻訳装置に関するもので、より詳細には語彙変更機能を有する機械翻訳装置に関するものである。

従来、ある言語を生成するための手段としては、ある概念を表す抽象構造(概念構造)から、概念構造中の品詞(名詞、動詞、形容詞、副詞など)間の関係を基に文層の単語列を決定していく。しかしながら自然な文章を生成するには、いわゆる文法上の品詞分けでは不十分であり、当該言語において適切な語彙変形を使用者が行う必要があった。

また上記の生成方法では、修飾關係があいまいになる場合があり、原文の概念構造の意味を忠実に表現する該文(文層の単語列)を必ずしも生成できなかった。

卷之三

本発明は、上述のように多種類に纏めてなされたものが、個々の繊維表面の荷電を失わないままで、しかも自然な繊維である状態を生成するために、繊維表面を有する繊維表面溶液を提供する手段を目的としてなされたものである。

卷之三

本発明は、上記目的を達成するため、(1)少なくとも、対象言語の原文を入力するための入力手段と、該入力手段により入力された原文を記憶する記憶手段と、言語解析に必要な知識を記憶している解析手段と、前記記憶手段に記憶された対象言語の原文を該記憶手段を用いて目的言語に翻訳する翻訳手段と、前記入力手段による入力された原文を該記憶手段により翻訳された翻訳結果を表示するための表示手段と、該文の文筋内の体言及び用言、該動詞(複合句)を検出する検出手段と、該検出手段の検出結果によって該文の語調を変更する語調変更手段とを有すること、更には、(2)前記語調変更手段において、語調変

立することである。図中、1.1 は形態素解析器、1.2 は発音、1.3 は語彙語形登録ファイル、1.4 は情報ファイル、1.5 は語彙文書解析器、1.6 は文法規則、1.7 は発音ルール群、1.8 は英語構造固定器、1.9 は情報ファイル群、2.0 は解析ルール群、2.1 は情報ファイル群、2.2 は変換器、2.3 は情報ファイル群、2.4 は構文生成器、2.5 は形態素生成器である。機械翻訳システム演習中の辞書 1.3 には、語形、活用形、品類、誤認などを記述されている。

入力された英字は、まず形態素解析部によって一つ一つの単語として認定され、語幹+助詞や数字+単語などの特別な文字符は名前としてまとめられる。

構文解析器 1.0 で入力英文の構造が解析される。文末から始めて右から左に、辞書登録名詞と文法ルールを適用するに従って文構造を決定していく。例えば、形容詞 + 動詞 + 名詞が副詞句と解釈される。

英語翻訳を終了した後は、係り受け関係等を理解するため動詞ノードを調べる本機能が

更対象文語間に拘束点を付与する付与手段を有すること、更には、(3)前級語語彙変換手段において、前級検出手段の検出結果だけでなく、該文側の語彙も反映した語彙変換手段を有することを特徴としたものである、すなわち、本発明では、該文の文節内の体言及び用言、終助詞(相違句)を検出し、その検出結果によって該文の語彙を変更する語彙変換手段と、依存関係にない修飾語係になり得る文節間に拘束点を付与する付与手段とを有することを特徴としたものである。これによって、あいまいな修飾關係を排除することができる。該文の概念構造に忠実で、かつ理解容易な該文を生成することができる。以下、本発明の実施例に基づいて説明する。

第1圖は、本発明による機械翻訳装置の一実施例を説明するための構成図で、ここでは英日機械翻訳装置の場合について説明する。以下、特にこれとわらぬ限り、英日機械翻訳装置に關して言及するが、本発明による方式は通常の制約が少ない文法体系を持つ言語（藤着語）に対して一般的に成

作られ、副説等の盛り充等が決まる、次に、共通条件の成立が調べられ、妥当な根拠が取られる。

前記木構造は変換部 2-2 で日本語の言い出しを
読みし語り、中國構造に翻訳される。構文生成部
ならでは文節単位で日本語を生成した後、形態素
生成部 2-3 で語彙の並びを決定したり、用語の活
用形を算定する。

実験部 2.2 からの出方は、以下のようないきなり標榜を希望している。

卷之三

1	勞 務 部
2	勞 務 部 辦 公 室
3	勞 務 部 辦 公 室 秘 書
4	勞 務 部 辦 公 室 秘 書 兼 公 關 經 理

このような従容關係から、説文を生成する。日本語は、文節内の語類は厳密に決まっているが、文節の間に關しては比較的の自由である（語類が文

尾、修飾文節は被修飾文節に先行する。など)としたがって以下の3文が生成されたとする。

①私は、私が以前見たことのあるきれいな女性に公園で会った。

②私は、きれいな私が以前見たことがある女性に公園で会った。

③私は、公園で私が以前見たことのあるきれいな女性に会った。

生成された表層文①の係受け関係は依存関係を実に表現している。一方、表層文③においては、複数の係受け関係が可能であり、誤った依存関係に解釈されやすい。例文は③においては、「きれいな」のは「私」と解釈されやすく、③では「公園で」「以前見たことのある」「女性」に

「会った」と解釈されやすい。この理由は、被修飾文節は、修飾可能な文節の内で最も近くの文節を修飾する傾向があるからである。

原文の概念構造の持つ意味を失わないまま、しかも自然で理解容易な表文を生成するためには、本文内のあいまいな修飾関係を排除する難易度

の高さが復雑だが、小さな範囲(内部)の係受けから広い範囲の文節の係受けへと、順に決定していく。接続詞をまたがって語順を変えることはない。

まず、連体修飾関係について説明する。連体修飾語(語)となる文節を次のように分類する。

連体修飾語となる文節の分類

(1) 体言密接グループ

形容詞(連体形) → 美い、大きい

連体副詞 → この、あるゆる、大変な、

形容動詞(連体形) → 駭かな、静かな、可憐な

(2) 体言非密接グループ

(2.1) 体言を含まない

(2.1.a) 動詞 (連体形)

→ 駆せる、走る、擦られる

(2.1.b) 運用修飾語+動詞(連体形)

→ 駆く走る、擦しく擦れる

(2.2) 体言を含む

(2.2.a) 専属形を含むもの

→ 密率がよい、息を切らした

機能が必要である。日本語では、最も複雑な文節は他の文節に先行させ、文節の位置に配置したほうがよい。これは、主語や他の文節と連語の間に複数の複雑な文節が置かれるとき、複雑な方に先行する文節の構文上の係受け関係がわかりにくくなり、かつ読み難くなるからである。

機能の複雑な文節を判定する方法は、当該文節内の階級類(相当句)を数えればよい。また連体修飾の場合は、修飾文節と被修飾文節との間に他の外意があると被修飾文節がその外意に係りやすくなり、誤った依存関係に解釈してしまう。両端に運用修飾の場合は、修飾文節と被修飾文節との間に他の用意があると運用修飾文節がその用意に係りやすくなる。

以上の点を考慮して、具体的な実現手順を示す。次に示す順序にしたがって語順を調整する。いずれも兄弟関係の語順に属する語形である。ここでいう兄弟関係とは、依存関係において、同じ文節を修飾する、同一レベルの文節の関係を指す。

原則としては可能な限り修飾文節は被修飾文節

(2.2.b) 体言を含む

→ フランスの、寒い、男性用の、
連体修飾関係における語順規則は以下の通りである。

(1) 体言密接グループ

それが係る体言の近くに置く。体言密接グループ内の語順は、文節長が短いものほど体言に近い。このグループ内には統点を含まない。文節の長さとしては、文字数や音節数などが考えられる。

(2) 体言非密接グループ

上記「体言密接グループ」内の語順を決定した後、その体言を含んだ文節を修飾する。

修飾される体言を含んだ文節から遠い順に

(2.2.a) (2.2.b) (2.2.c) (2.2.d) → (体言を含む文節)
体言を含む 体言を含まない

の順に並べる。同一グループ内では文節長の短いものが体言文節に近い。

あいまいな係受け関係を避けるために、隣接する文節間に以下の統点規則を設ける。

◎統点を付与しない

「体言を含む文節」「体言を含まない文節」

例 効率がよく遠く走る→車

「体言を含まない文節」「体言を含まない文節」

例 遠く走る車→車

◎統点を付与する

「体言を含む文節」「体言を含む文節」

例 効率がよくフランスの→車

「体言を含まない文節」「体言を含む文節」

例 遠く走る→フランスの→車

次に、適用修飾関係について説明する。適用修飾関係(語)となる文節を次のように分類する。

適用修飾語となる文節の分類

(1) 用言密接グループ

副詞 ……ふと、さらに、いくぶん

形容詞(適用形) ……赤く、大きく、遠く

形容動詞(適用形) ……春かに、素直に

(2) 用言非密接グループ

(2.1) 用言を含まない

体言+格助詞(相当句)→イシタで、東京駅から

(2.2) 用言を含む

(2.2.1) 格要素を含むもの

…行儀がよく、急を切らして、

(2.2.2) 数詞(適用形)+接続助詞

…軽かしがれて、食べながら、

(2.2.3) 用言を含む名詞句+格助詞(相当句)

…赤いインクで、遠く走る車によって、
遠体修飾関係における類属類別は以下の通りである。

(3) 用言密接グループ

それが保有用言の近くに置く。用言密接グループでの類属は、「程度を表わす副詞」が、それが修飾する用言に最も近く、その他は文節長が短いものほどそれが修飾する用言に近い。このグループには統点を付与しない。

(4) 用言非密接グループ

「用言密接グループ」内の類属を決定した後、その用言を含んだ文節を参照する。

修飾される用言を含んだ文節から遙い順に

という類属になる。

次に、文節「女性における適用修飾語は、「きれいな」「私が以前に見たことのある」の2つである。「きれいな」は適用修飾語(語)、「私が以前見たことのある」は格要素を含むので、用言非密接グループである。よって類属は、「私が以前見たことのある」「きれいな」→「女性による」となる。

最後に、文節「会った」への適用修飾語は、「私は」「私が以前見たことのあるきれいな女性だ」「公園で」の3つである(主語も適用修飾語の一類と考える)。

「私は」と「公園で」は、用言非密接グループの(1.1)である。「私が以前見たことのあるきれいな女性だ」は、用言非密接グループの(2.2.2)である。したがって類属は、

「私が以前見たことのあるきれいな女性だ」「公園で」「私は」→「会った」となる。

最終的に生成される表層文は、

(2.2.3) (2.2.3.2) → (2.1) → (用言を含む文節)

…用言を含む ……用言を含まない

の順に並べる。両グループ内では文節長が短いものほど、用言文節に近い。

あいまいな語り受け関係を避けるために、隣接する文節間に以下の統点規則を設ける。

◎統点を付与しない

「用言を含む文節」「用言を含まない文節」

例 行儀がよく車で→食べる

「用言を含まない文節」「用言を含まない文節」

例 東京駅から車で→来る

◎統点を付与する

「用言を含む文節」「用言を含む文節」

例 行儀がよく、數詞しながら→食べる

「用言を含まない文節」「用言を含む文節」

例 電話によって、食べながら→連絡する

先に示した依存関係に本発明の語彙変換手段を施した例を示す。最初に、最も下位の依存関係の語彙が始められる。その結果、

「私が」「以前」→「見たことのある」

「私がお前見なごとのあるそれがいな友様に会議で
私は会った。」
となる。

第2類は、以上に説明したような語彙整形手順のフローチャートである。以下、各ステップに従って順に説明する。

step.1 まず、処理対象を移動する。

step.2 1. 当該文節の文節グループ判定を行う。

step.3 1. 兄弟關係の文節があるかどうか判断する。
2. step.2の語彙規則に従って決定する。

step.4 読点規則に従って読点付与を行う。

step.5 1. 魚子關係があるかどうか判断する。魚子
關係がなければ終了し、あればstep.1に戻る。

本発明による文節の分離手順、すなわち文節中の体言・用言の移動は容易に実現できる。
まず第1段のように述語中の該語情報に体言・用言の移動を記入しておく。また格要素の有無も同時に付与しておく。

文節生成時には、これらの情報を受け継いでいくことにより、生成された文節に対しても判定す

ることができる。

第1表 読点規則

見出し語	状語	格助詞数	体言	用言	音節数
efficient	有効な	0	無	素	0
	効果がある	1	有	表	0

次に源文側の情報を探査した語彙変更手順について説明する。

並列(等位)構造

並列構造の場合、並列構造の範囲内の該語は原文に基づき、読点は付与しない。例えば、

“beautiful brother-in-law and John”

→「美しい(義理の兄弟とジョン)」

→「(美しい義理の兄弟)とジョン」

の場合、英文側でも複数の修飾関係になり得るので、原文のあいまい性を原文においても保存する。
したがって上記名類句の該は

「美しい義理の兄弟とジョン」

となる。

共通補語を持つ並列の場合も同様に、該は原文に従う。

に記述する。

“Tough he studied hard, he ~~perfectly~~ failed in his exams.”

「彼は一生懸命勉強したけれども、玉砕に丸く終は彼の試験に落ちた。」

複合名詞

英語側でも複数形でありながら、和文側では、複体修飾のために「の」を挿入している場合には、いかなる場合でも名詞連続の範囲内に修飾成分は挿入しない。

慣用的な表現

慣用的表現では述語の直前に述用修飾成分を置くと奇異な日本語になる場合があるので、このような場合には、慣用的な表現の範囲の直前に修飾成分を配置する。

“He ~~enjoys~~ makesakes fun of old man.”

「彼は老人をときどき馬鹿にする。」

「馬鹿にときどきする」とは言わないと。

本発明は以上の実施例に規定されるものではなく、該文の文節内の体言及び用言、格助詞(格助

“He ~~saw~~ ~~sawed~~ wood and ~~sawed~~ his son.”

「彼は彼の息子を見つめ育んで、そして死る。」
動動詞の並列の場合も同様である。

“He ~~saw~~ and ~~sawed~~ do it well”

「かれはそれを上手にできる。そしてしなければならぬ。」

新記の「でき」のように述用形に活用せず、終止形のままにする方法もある。

「彼はそれを上手にできる。そしてしなければならぬ。」

英文側での特殊な出現位置

英文側において文節部にあり、和文側で述用修飾となるものは、和文側でも先頭に読点を添えて挿入する。

“He ~~was~~ absent from school yesterday.”

「彼は学校を休んだ。」

“Yesterday he ~~was~~ ~~was~~ absent from school.”

「昨日、彼は学校を休んだ。」

また、英文側で副詞の出現をコンマで区切られていたら、和文側では該点を添えてその文の先頭

等)を検出し、その検出結果によって原文の語彙を変更する語彙変換手段と、依存関係にない語彙関係になり得る文節間に統度を付与する語序手段とを有し、これによつて、あいまいな語彙関係を排除し、原文の概念構造に忠実で、かつ理解容易な新文を生成することが可能な語彙変換機能をもつ機械翻訳システムを提供することに本発明がある。

発明の実用

以上の説明から明らかなように、本発明によると、以下のようないくつかの効果がある。

(1) 語彙現象に対する効果：原文におけるあいまいな語彙関係を排除し、原文の概念構造に忠実で、かつ理解容易な新文を生成することができる。

(2) 語彙現象に対する効果：原文側の語彙のある程度反映しているので、原文と新文との間での印象があまり變化せず、したがつて読者側で整合しやすくなる。

4. 図版の簡単な説明

第1図は、本発明による機械翻訳装置の一実施

例を説明するための構成図。第2図は、該機械形手順のフローチャートである。

1.1…形態素解析部、1.2…統度、1.3…語彙解析部、1.4…情報ファイル1、1.5…構文解析部、1.6…文法規則、1.7…解析ルール1、1.8…英語構造用定数、1.9…情報ファイル2、2.0…解析ルール2、2.1…情報ファイル3、2.2…構文生成部、2.3…形態素生成部。

特許出願人 株式会社 日立製作所
代理人 高野義雄
(ほか1名)

